

Gliwice, 3 marca 2017 roku

dr hab. inż. Krzysztof Cyran, prof. nzw. w Pol. Śl.
Instytut Informatyki
Wydział Automatyki, Elektroniki i Informatyki
Politechnika Śląska

RECENZJA ROZPRAWY DOKTORSKIEJ

Tytuł rozprawy: Dwuetapowa metoda eksploracji danych pozyskiwanych z obrazów cyfrowych

Autor rozprawy: mgr inż. Anna Wójcicka

Promotor rozprawy: prof. dr hab. inż. Zygmunt Wróbel

Dziedzina: nauki techniczne

Dyscyplina: informatyka

Cel, zakres i tezy rozprawy

Cel i zakres pracy jak również hipotezy badawcze służące do sformułowania tezy, wskazane są w podrozdziale 1.1. Na stronach 7-9 podano szczegółowo zagadnienia, które Autorka zamierzała zbadać wraz z przyjętą przez nią metodologią. W rozprawie, na podstawie swych badań wstępnych z pozyskiwania wiedzy z obrazów wykorzystywanych w dziedzinie inżynierii materiałowej (Friction Stir Welding) oraz w medycynie, Autorka sformułowała cztery hipotezy badawcze (hipotezę dotyczącą możliwości wyboru metod analizy obrazów najbardziej adekwatnych dla potrzeb późniejszej eksploracji wiedzy z otrzymanych z tej analizy wyników, hipotezę w pewien sposób odwrotną, wskazującą na możliwości wyboru metod eksploracji danych, najbardziej odpowiednich do wydobywania wiedzy z wyników uprzednio użytych metod analizy obrazów oraz hipotezę możliwości połączenia obu podejść w dwuetapowej metodzie odkrywania wiedzy zawartej w obrazach cyfrowych, co z jednej strony, zdaniem Autorki, ma zapewnić częściowe zautomatyzowanie całego procesu, z drugiej zaś umożliwić ingerencję ekspertów dziedzinowych w proces pozyskiwania wiedzy – możliwość tej ingerencji stanowi treść ostatniej, czwartej hipotezy).

W oparciu o tak podane hipotezy badawcze, Autorka stawia tezę rozprawy, iż opracowanie dwuetapowej metody pozyskiwania wiedzy z obrazów cyfrowych polegającej na ekstrakcji z obrazów cech ilościowych i jakościowych, a następnie na eksploracyjnej analizie danych, stanowić będzie znaczący i użyteczny krok przy tworzeniu regułowych baz wiedzy przeznaczonych dla systemów wspomagania decyzji. Autorka, starała się udowodnić tak sformułowaną tezę przez opracowanie systemu informatycznego, działającego w oparciu o proponowaną dwuetapową metodę, w którym połączenie metod przetwarzania obrazów oraz eksploracji danych wykorzystuje formalizm teorii zbiorów przybliżonych, zaś system ma za zadanie użyć uzyskane reguły dla celów klasyfikacyjnych.

Należy zauważyć, że sama dwuetapowość podejścia wykorzystywanego przy wydobywaniu wiedzy użytecznej z obrazów cyfrowych na użytek klasyfikacji nie jest w żaden sposób

podejściem nowatorskim ani oryginalnym – jest znanym od dawna połączeniem ekstrakcji z obrazów cech charakterystycznych z metodami uczenia klasyfikatorów w przestrzeni tychże cech. Dwuetapowość ta jest prostą konsekwencją poszukiwania metod radzenia sobie z tzw. „przekleństwem wymiarowości”, problemem szczególnie wyraźnym przy klasyfikacji obrazów cyfrowych, które z punktu widzenia problemów uczenia maszynowego nie są przecież tworami 2-wymiarowymi, lecz wektorami $(n \times m)$ -wymiarowymi (co przy typowych wartościach szerokości n oraz wysokości m obrazu, lokuje problem w przestrzeniach o setkach tysięcy lub milionach wymiarów – z wszystkimi tego faktu konsekwencjami).

Na miano oryginalnego zasługuje za to podejście, starające się połączyć oba te etapy w spójną metodę, charakteryzującą się z jednej strony maksymalizacją stopnia jej zautomatyzowania, z drugiej zaś uwzględnieniem możliwości ingerencji ekspertów dziedzinowych, w sytuacjach gdy wiedza dziedzinowa zawiera elementy szczególne, których wykorzystanie zapewnia lepsze wyniki, niż bazowanie wyłącznie na metodach ogólnych.

Ocena zawartości pracy

Praca składa się z dziewięciu rozdziałów, w tym oznaczonego jako rozdział 1 wprowadzenia, oraz stanowiącego rozdział 9 podsumowania. Całość rozpoczyna się spisem treści, a kończy spisem rysunków, tabel oraz liczącym 111 pozycji wykazem literatury.

We wprowadzeniu (oznaczonym w pracy jako rozdział 1), oprócz przedstawienia celu i tez rozprawy, których ocenę przedstawiłem w poprzedniej części, podano w podrozdziale 1.2 strukturę pracy, co w zamierzeniu powinno było ułatwić poruszanie się po przedstawionej rozprawie, jednakże posłużono się do tego wadliwą, bo przesuniętą o 1, numeracją rozdziałów w stosunku do numeracji rzeczywiście zastosowanej. Chodzi o to, że w pracy oznaczono wprowadzenie jako rozdział 1, natomiast w opisie struktury pracy jako rozdział 1 potraktowano rozdział dotyczący motywacji i koncepcji proponowanego rozwiązania (w rzeczywistości jest to rozdział oznaczony numerem 2). To przesunięcie stosowano w opisie struktury konsekwentnie, kończąc na podsumowaniu, które w związku z tym zostało przedstawione jako rozdział ósmy, podczas gdy w istocie w pracy jest oznaczone numerem 9. Samo zamieszczenie we wprowadzeniu podrozdziału dotyczącego struktury pracy uważam za cenne, gdyż w ten sposób Autorka zwięźle pokazuje, co czeka czytelnika w dalszej części rozprawy. Jednakże, ze względu na wspomnianą wyżej pomyłkę, korzyści z wykorzystania tego podrozdziału są okupione koniecznością „odkrycia” przez czytelnika reguły przesunięcia numeracji, a następnie jej konsekwentnego stosowania polegającego na dodawaniu 1 do numeru rozdziału opisanego w strukturze pracy, po to by mógł on trafić do właściwej części rozprawy. W niniejszej recenzji, przy referencjach do wybranych fragmentów pracy, stosuję faktyczną numerację rozdziałów rozprawy, a nie tę podaną w podrozdziale 1.2 Struktura pracy.

W rozdziale drugim przedstawiono powody, dla których Autorkę podjęła się badań zmierzających do stworzenia dwuetapowego systemu eksploracji wiedzy z obrazów cyfrowych. Wskazała w nim również formalizmy, które stanowią element sprzęgający dwa etapy proponowanej metody. Centralne miejsce zajmuje tutaj system informatyczny Pawlaka, który ma być wynikiem ekstrakcji cech i bazą do dalszej eksploracji wiedzy metodami znanymi z teorii zbiorów przybliżonych (z wszystkimi tego konsekwencjami – niewątpliwymi zaletami takimi jak naturalna reprezentacja tablicy decyzyjnej, dla której opracowano aparat formalnej weryfikacji spójności wiedzy jak również jej eksploracji i budowania reguł decyzyjnych, oraz wadami, występującymi przynajmniej w klasycznej

wersji teorii zbiorów przybliżonych, m.in. koniecznością – często sztucznej – dyskretyzacji atrybutów ciągłych).

Rozdział trzeci dotyczy przeglądu metod analizy i przetwarzania obrazów. Cenną jest próba pogrupowania wielu dostępnych metod w usystematyzowaną całość (skrótowo przedstawiona na rysunku 3.1) – jednakże do użytego nazewnictwa jednej z głównych kategorii mam zastrzeżenia (co podaję, wraz z uzasadnieniem w sekcji dotyczącej niedociągnięć rozprawy). Następnie, konsekwentnie do dokonanej kategoryzacji, następuje opis poszczególnych metod, w zdecydowanej większości jasny i poprawny, jednakże, w paru miejscach i tu, nie ustrzeżono się nieścisłości (opisanych w sekcji dotyczącej wad pracy).

W rozdziale czwartym Autorka dokonuje przeglądu metod służących eksploracji danych, pod kątem ich zastosowania w drugiej fazie swej strategii. Skupia się przede wszystkim na metodach regułowych, jako najlepiej przystających do reprezentacji wiedzy w postaci systemu informacyjnego Pawlaka, właściwie uwypuklając fundamentalną różnicę pomiędzy regułami decyzyjnymi, w których z góry znany jest atrybut występujący w konkluzji, a regułami asocjacyjnymi, dla których atrybuty występujące w konkluzji odkrywane są dopiero na etapie generowania reguł. Słusznie podkreśla Autorka wagę obu typów reguł w proponowanym przez siebie podejściu służącym do pozyskiwania wiedzy z (przede wszystkim medycznych) obrazów cyfrowych.

Rozdział piąty to krótki przewodnik po teorii zbiorów przybliżonych. W odczuciu recenzenta jest to najslabszy rozdział rozprawy – można się co prawda zgodzić, że dla kompletności dzieła, przytoczenie podstawowych pojęć teorii zbiorów przybliżonych może być przydatne, ale ograniczenie się w opisie wyłącznie do pojęć klasycznej teorii zbiorów przybliżonych w czasach, gdy znane są jej uogólnienia lub modyfikacje (niektóre bardzo dobrze dostosowane do specyfiki obrazów jako danych wejściowych) uważam za niewystarczające. Liczę tu na dyskusję w trakcie obrony.

W rozdziale szóstym Autorka dokonała przeglądu metod generowania zbiorów reguł. Opisała problem generowania reguł decyzyjnych i asocjacyjnych. W ramach reguł decyzyjnych omówiła bardziej szczegółowo metody oparte o teorię zbiorów przybliżonych (w tym prawidłowo zwróciła uwagę na znaczenie uogólnionego reduktu w generowaniu reguł decyzyjnych i omówiła skrótowo algorytmy generowania reguł, w tym m.in.: LEM1, LEM2, czy MODLEM), a następnie skupiła się na charakterystyce innych podejść, odnosząc się do metody selektywnego pokrywania (i opartych na niej algorytmów generowania reguł AQ oraz CN2). W ramach algorytmów generujących reguły asocjacyjne, większą uwagę Autorka poświęciła algorytmowi Apriori, znacznie szybszemu algorytmowi FP-Growth, oraz algorytmowi odkrywania asocjacji GSP. Rozdział ten, choć operuje na wysokim poziomie abstrakcji i unika wchodzenia w szczegóły, jest napisany poprawnie i zawiera właściwie wyselekcjonowane informacje, ora krótką, w ocenie recenzenta, sensowną dyskusję dotyczącą znaczenia reguł asocjacyjnych w odkrywaniu wiedzy. Niestety rozdział zawiera też (na szczęście tylko sporadycznie) niepotrzebne powtórzenia (wskazane w części recenzji, która dotyczy wad rozprawy).

System eksploracji wiedzy z obrazów medycznych, zbudowany na kanwie proponowanej przez Doktorantkę dwuetapowej metody, to treść rozdziału siódmego. Autorka opisuje w nim najpierw część dotyczącą ekstrakcji cech charakterystycznych obrazów (używając jako przykładu obrazów dermatoskopowych nowotworu skóry, udostępnionych jej do badań, jak podaje Autorka, przez AGH w Krakowie), a następnie część dotyczącą wydobywania wiedzy

(z wykorzystaniem metodologii teorii zbiorów przybliżonych). Całość kończą uwagi implementacyjne dotyczące m.in. użytego języka programowania oraz wykorzystanych bibliotek. W rozdziale, niestety nie ma wzmianki o kwestiach etycznych związanych z przetwarzaniem obrazów medycznych – można się spodziewać, że Doktorantka pracowała na danych anonimowych – jednakże, winna to jasno wskazać w rozprawie.

W rozdziale ósmym przedstawiono wyniki badań eksperymentalnych przeprowadzonych z użyciem opisanego wcześniej oprogramowania Doktorantki. Jako materiał do badań wykorzystwała serię zdjęć stomatologicznych wykonanych za pomocą tomografu optycznego oraz obrazy spoin wykonanych w technologii FSW. Bazując na zaproponowanej przez siebie metodzie i posiłkując się stworzonym przez siebie oprogramowaniem oraz wiedzą ekspercką ekspertów dziedzinowych, Autorka przeprowadza eksperymenty prowadzące do odkrycia reguł decyzyjnych. Jest to w opinii recenzenta najciekawszy fragment rozprawy, gdyż widać w nim praktyczną przydatność zaproponowanego przez Doktorantkę podejścia, jednakże i w tej części nie wszystko zostało wystarczająco uzasadnione. I tak np. na końcu str. 93 Autorka stwierdza, że „rozmiar maski określono na podstawie wielkości występujących szumów i artefaktów”. Nie wyjaśnia jednak jak tego dokonano – w jaki sposób szумы i artefakty o określonej (nigdzie nie podanej) wielkości wpłynęły na wybór maski o rozmiarze 7x7. Recenzent liczy, że obrona pracy będzie dobrą okazją do wyjaśnienia przyjętej w tym zakresie metodologii. Podobnie, na str. 94 nie przeprowadzono żadnej dyskusji nad parametrem k (według słów autorki, zależnym od problemu) koniecznym do wyznaczenia wartości progu w metodzie binaryzacji Sauvola. Również nie jest jasnym z rozprawy w jaki sposób, dla centralnego fragmentu 100x100 pikseli „odczytana została grubość szkliwa” – należy domniemywać, że grubość szkliwa w danym punkcie odpowiada jasności odpowiedniego piksela, jednakże nie zostało to jasno napisane, a dla osób nie znających sposobu interpretacji zdjęć otrzymanych z tomografu optycznego, taka interpretacja wcale nie jest jedyną z możliwych. Recenzent liczy na odniesienie się na obronie pracy co do prawdziwości (bądź nie) tego domniemania. Również chciałby uzyskać więcej szczegółów na temat walidacji krzyżowej reguł decyzyjnych, gdyż poza przytoczonym wynikiem tej walidacji (70%) nie przeprowadzono żadnej dyskusji co ten wynik oznacza dla praktycznego zastosowania klinicznego uzyskanego zbioru reguł. Podobnie nie poddano dyskusji problemowi walidacji krzyżowej reguł (78%) otrzymanych za pomocą dwuetapowej metody Doktorantki, w wyniku czego otrzymała system ekspertowy określający (niestety dyskretyzowaną) siłę spoiwa wykonanego metodą FSW.

Całość rozprawy znajduje podsumowanie w rozdziale dziewiątym.

Opinia o pracy

Kandydatka w dość szerokim zakresie opanowała warsztat metod przetwarzania obrazów i przetestowała dużą ich liczbę na danych medycznych (stomatologicznych) uzyskanych za pomocą tomografu optycznego oraz obrazach spoin wykonanych metodą FSW. Ma również znajomość problematyki dotyczącej eksploracji wiedzy, zwłaszcza z wykorzystaniem teorii zbiorów przybliżonych, oraz potrafiła twórczo połączyć te dziedziny w proponowanej dwuetapowej metodologii, co stanowi o elemencie oryginalności pracy. Cennym elementem jest również wykorzystanie przez doktorantkę opracowanej metodologii do stworzenia praktycznego systemu informacyjnego służącego eksploracji danych z obrazów (przede wszystkim medycznych). Dlatego krytyczne uwagi, zarówno te już wzmiankowane, jak i te wymienione poniżej, nie mają w opinii recenzenta znaczenia pierwszorzędного, niektóre z nich są dyskusyjne i zapewne doczekają się polemiki doktorantki w trakcie obrony. Są jednak

i takie, których z pewnością dało się uniknąć, co byłoby z korzyścią dla wartości zaprezentowanego w rozprawie dorobku i pozwoliłoby czytającemu mieć lepsze zdanie o całości pracy.

Wady i słabe strony rozprawy oraz krytyczne uwagi szczegółowe

Do wad pracy można zaliczyć:

- W odczuciu recenzenta niewystarczające powoływanie się na *state-of-the-art* (dość dobrym tego przykładem jest konstatacja Doktorantki zamieszczona na str. 6, iż „studia literaturowe przeprowadzone przez autorkę wykazały jednak, że podejście spójne, łączące te dwie dziedziny są niezwykle rzadko stosowane”, po którym to, dość stanowczym twierdzeniu, brakuje całkowicie powołania się na jakąkolwiek literaturę przedmiotu – jasnym jest, że celem Autorki jest tutaj wykazanie braku badań o wskazanych cechach, jednakże wypadałoby przywołać choć publikacje opisujące przykłady badań ograniczonych do jednej lub drugiej dziedziny).
- Dyskusyjne jest nazywanie grupy metod przetwarzania obrazów zawierającej m.in. takie operacje jak obrót, przesunięcie, zmiana histogramu itp., przekształceniami w dziedzinie czasu – w prostym odróżnieniu od operacji w dziedzinie częstotliwości (np. transformaty Fouriera). O ile takie nazewnictwo (dziedzina czasu vs. dziedzina częstotliwości) nie budzi zastrzeżeń dla analizy sygnałów jednowymiarowych (powstałych właśnie w dziedzinie czasu), o tyle przejście do drugiego wymiaru występującego w obrazach, powinno, w opinii recenzenta, uwzględniać podstawową semantykę pojęcia „czas”, którą najwyraźniej zaniedbuje Autorka. W przypadku dwuwymiarowych obrazów w wyniku zastosowania transformat, dziedzina częstotliwości staje się ściśle biorąc dziedziną częstotliwości przestrzennych – jednak pominięcie przymiotnika „przestrzenny” jest jak najbardziej dopuszczalne i mówienie po prostu o metodach operujących w dziedzinie częstotliwości nie wprowadza niejasności. Jednak, w przypadku danych wejściowych nie poddanych transformacjom częstotliwościowym, mówienie o dziedzinie „czasu” ma sens jedynie wówczas gdy dane faktycznie mają naturę czasową, są przebiegami czasowymi. W przypadku obrazów, operacje, które Autorka grupuje na rys. 3.1 (także, konsekwentnie na rys. 3.26 i 3.27) jako działające w dziedzinie czasu, w żaden sposób do czasu się nie odnoszą – i dlatego nazywanie ich przekształceniami w dziedzinie czasu budzi moje zastrzeżenia.
- Opis słowny operacji binaryzacji warunkowej (z histerezą) nie odpowiada opisowi formalnemu przedstawionemu we wzorze (3.10). Poza tym nie podano jak rozumiane jest sąsiedztwo w przytoczonym wyjaśnieniu do tego wzoru: „s – wartość sąsiadujących punktów”.
- Przytaczanie w rozprawie doktorskiej wiadomości elementarnych – np. na str. 32 Autorka nie tylko wyjaśnia czym jest mediana, ale również jako ilustrację przytacza przykładową 9-elementową maskę (3x3), z której następnie tworzy ciąg uporządkowanych wartości (czyni to zresztą błędnie, bowiem jej ciąg wynikowy ma 8 elementów a nie 9), po czym wskazuje wartość mediany tego ciągu. Zwracam uwagę, że to nie pomyłka polegająca na błędnym pominięciu liczby 45 jest tutaj podstawowym przedmiotem krytyki, lecz sam fakt zamieszczania w pracy naukowej tak trywialnych ilustracji.

- Niedokładne przedstawianie opisów bibliograficznych (np. w pozycji [1] brak jest podania nazwy czasopisma, w którym zamieszczono artykuł – przeszukiwanie internetowe pokazuje, że chodzi o publikację zamieszczoną w Studia Informatica, jednakże w pracy nie zamieszczono tej informacji).
- Niestety dość liczne błędy pisarskie (m.in. na str. 6 jest „ceny materiał” zamiast „cenny materiał”, na str. 10 jest „regulową reprezentacje” zamiast „regulową reprezentację”, na str. 26 jest „polor czarny” zamiast „kolor czarny” oraz „wartość odpowiadającej” zamiast „wartość odpowiadającą”, na str. 31 jest „kontem” zamiast „kątem”, na str. 37 jest „z punktem centralnych” zamiast „z punktem centralnym”, na str. 50 jest „owych definicje” zamiast „owych definicji”, na str. 54 jest „niezależne metoda” zamiast „niezależna metoda”, na str. 57 jest „mogą ona” zamiast „mogą one”, na str. 66 jest „macierzy nieodróżnialności” zamiast „macierzy rozróżnialności”, na str. 69 jest „tablice decyzyjną” zamiast „tablicę decyzyjną”, na stronie 86 w tytule sekcji 7.1.3 jest „Prezentację wyników 1-go etapu” zamiast „Prezentacja wyników 1-go etapu” oraz aż trzykrotnie jest „Colur” zamiast „Colour” lub „Color”, na str. 94 jest „na przefiltrowanych obrazy” zamiast „na przefiltrowanych obrazach” – to stosunkowo pokaźne wyliczenie wskazuje niektóre tylko z błędów pisarskich – zawsze mogą takie błędy się zdarzyć, jednak ich ilość powinna być znacząco mniejsza).
- Niedokończone wypowiedzi, np. na str. 37 znajduje się następujące „zdanie”: „W przypadku dyskretnej reprezentacji obrazu, z wyróżnionym jednym punktem (tzw. punktem centralnym).”.
- Powtórzenia tekstu występujące w rozdziale 6 – akapit rozpoczynający podrozdział 6.3, został zamieszczony również jako ostatni podrozdziału 6.2.
- Sygnalizowana wcześniej niespójność pomiędzy faktyczną numeracją rozdziałów a numeracją występującą w podrozdziale 1.2 Struktura pracy.

Wniosek o dopuszczenie rozprawy doktorskiej do publicznej obrony

Podsumowując, stwierdzam, że rozprawa Pani mgr inż. Anny Wójcickiej, pomimo podanych wyżej uwag krytycznych, nosi znamiona oryginalności, świadczy o ogólnej wiedzy Doktorantki w zakresie eksploracji wiedzy z obrazów cyfrowych oraz wykazuje umiejętność Autorki samodzielnego prowadzenia pracy naukowej, w związku z czym spełnia wymagania odnośnej Ustawy. Biorąc to pod uwagę, wnioskuję o dopuszczenie jej do publicznej obrony.

